



# Why and Wherefore of Increased Scientific Collaboration

## Citation

Freeman, Richard B., Ina Ganguli, and Raviv Murciano-Goroff. 2015. Why and Wherefore of Increased Scientific Collaboration. Working paper.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:34305979>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

## **Why and Wherefore of Increased Scientific Collaboration<sup>1</sup>**

Richard B. Freeman (Harvard University and NBER)  
Ina Ganguli (SITE, Stockholm School of Economics)  
Raviv Murciano-Goroff (Stanford University)

August 28, 2014

JEL: F20, J24, O30

### **ABSTRACT**

This paper examines international and domestic collaborations using data from an original survey of corresponding authors and Web of Science data of articles that had at least one US coauthor in the fields of Particle and Field Physics, Nanoscience and Nanotechnology, and Biotechnology and Applied Microbiology. The data allow us to investigate the connections among coauthors and the views of corresponding authors about the collaboration. We have four main findings. First, we find that US collaborations have increased across US cities as well as across international borders, with the nature of collaborations across cities resembling that across countries. Second, face-to-face meetings are important in collaborations: most collaborators first met working in the same institution and communicate often through meetings with coauthors from distant locations. Third, the main reason for most collaborations is to combine the specialized knowledge and skills of coauthors, but there are substantial differences in the mode of collaborations between small lab-based science and big science, where international collaborations are more prevalent. Fourth, for biotech, we find that citations to international papers are higher compared to papers with domestic collaborators only, but not for the other two fields. Moreover, in all three fields, papers with the same number of coauthors had lower citations if they were international collaborations. Overall, our findings suggest that all collaborations are best viewed from a framework of collaborations across space broadly, rather than in terms of international as opposed to domestic collaborative activity.

---

<sup>1</sup> We appreciate assistance with the survey from John Trumbour and input from Jennifer Amadeo-Holl, Paula Stephan and Andrew Wang. We received helpful comments from Adam Jaffe, Ben Jones, Manuel Trajtenberg and participants at the NBER “The Changing Frontier: Rethinking Science and Innovation Policy” Conferences. This research was supported in part by the National Science Foundation’s National Nanotechnology Initiative, Award 0531146.

Scientists increasingly collaborate on research with other scientists, producing an upward trend in the numbers of authors on a paper (Jones, Wuchty and Uzzi, 2008; Wuchty, Jones and Uzzi, 2007; Adams, Black, Clemmons, and Stephan, 2005). Papers with larger numbers of authors garner more citations and are more likely to be published in journals with high impact factors than papers with fewer authors (Lawani, 1986, Katz and Hicks, 1997; deB. Beaver, 2004, Wuchty et al., 2007; Freeman and Huang, 2014), which seems to justify increased collaborations in terms of scientific productivity. The trend in coauthorship extends across country lines, with a larger proportion of papers coauthored by scientists from different countries (Indicators, 2012; Adams, 2013). In the United States and other advanced economies, the proportion of papers with international coauthors increased from the 1990s through the 2010s, while the proportion of papers with domestic coauthors stabilized. In emerging economies, where collaboration has not yet reached the proportions in the US and other advanced countries, the share of papers with domestic collaborations and the share with international collaborations have both increased.

The spread of scientific workers and research and development activity around the world (Freeman, 2010) has facilitated the increase in international collaborations. The growing number of science and engineering PhDs in developing countries, some of whom are international students and post-docs returning to their country of origins (Scellato, Franzoni, and Stephan, 2012) has expanded the supply of potential collaborators outside the North American and Western European research centers. A rising trend in government and industry R&D spending in developing countries and grant policies by the European Union and other countries favor international cooperation. At the same time, the lower cost of travel and communication has reduced the cost of collaborating with persons across geographic locales (Agrawal and Goldfarb, 2008; Catalini, Fons-Rosen and Gaulé, 2014). The increased presence of China in scientific research, exemplified by China's move from a modest producer of scientific papers to number two in scientific publications after the US, has been associated with huge increases in collaborations between Chinese scientists and those in other countries.<sup>2</sup>

Finally, the location of scientific equipment and materials, such as CERN's Giant Hadron Collider, huge telescopes in particular areas, or geological or climatological data available only in special localities, have also increased collaborations. The US was not a prime funder for

---

<sup>2</sup> Science and Engineering Indicators 2013, Appendix table 5-27 gives scientific papers for the top five countries in 2009: US 208,601; China 74,019; Japan 49,627; UK 45,649; and Germany 45,002.

CERN, but Americans are the largest group of scientists and engineers working at CERN. China eschewed joining the CERN initiative as an associate member state, but many China-born scientists and engineers work at CERN as members of research teams from other countries.

How successful are collaborations across country lines and across locations in the same country? How do collaborators meet and develop successful research projects? What are the main advantages and challenges in collaborative research?

To answer these questions, we combine data from a 2012 survey that we conducted of corresponding authors on collaborations *with at least one US coauthor* with bibliometric data from Web of Science (WoS) (Thomson Reuters, 2012) in three growing fields – Particle and Field Physics, Nanoscience and Nanotechnology, and Biotechnology and Applied Microbiology. The survey data allow us to investigate the connections among coauthors in collaborations and the views of corresponding authors about collaborations. The WoS data allows us to examine patterns of collaborations over time and to compare patterns found in our fields to those found in scientific publications broadly. To determine whether borders or space are the primary factor that affects the nature and impact of collaborations, we contrast collaborations across locations in the US, collaborations in the same city in the US, and collaborations with international researchers.

We find that US collaborations increased across US cities as well as internationally and that scientists involved in these collaborations and those who collaborate in the same locale report broad similarities in their experiences. Most collaborators first met while working in the same institution. Most say that face-to-face meetings are important in communicating with coauthors across distances. And most say that specialized knowledge and skills of coauthors drive their collaborations. We find that international collaborations have a statistically significantly higher citation rate than domestic collaborations only in biotech, a modestly higher citation rate in particle physics, but a lower rate in nanotech. Because international collaborations have a greater numbers of authors than other collaborations, once we account for the number of coauthors on papers, the higher citation rate for biotech and particle physics international collaborations also disappear. Our results suggest that the benefits to international collaboration in terms of citations depend on the scientific field in question, rather from any “international magic” operating on collaborations with the same number of researchers. By limiting our sample to papers with at least one US-based author, however, we exclude the

possibility that international collaborations greatly benefit researchers in countries with smaller research communities by linking them to experts outside their country, US aside.

## **1. The Growing Trend of International Collaboration**

We analyze data from corresponding authors and articles in which researchers collaborate in Particle and Field Physics, Nanoscience and Nanotechnology, and Biotechnology and Applied Microbiology. These three fields cover a wide span of scientific activity, with different research tools and methodologies.

Particle physics has a theoretical part and an empirical part. Leading edge empirical research requires massive investments in accelerators and colliders, of which the Large Hadron Collider is the most striking. Europe's decision to fund the Hadron Collider while the US's rejection to build a large collider in Texas shifted the geographic locus of empirical research from the US to Europe and arguably spurred the greater growth of string theory (which does not need direct access to the Collider) in the US than in Europe. Particle physics is the most mathematically and theoretically sophisticated of the sciences we study, where path-breaking mathematical analysis guides empirical work, and where the massive equipment exemplifies big science.

Nanotechnology is a general interdisciplinary applied technology, where engineers often collaborate with material scientists. The electron microscope is a pivotal research tool. The US made sizable investments in nanotechnology beginning at the turn of the 21<sup>st</sup> century, when President Clinton called for greater investment in nano-related science and technology. This led to the 21st Century Nanotechnology Research and Development Act that President Bush signed in 2003. Other countries undertook similar initiatives in the same period.

Biotechnology is lab-based, in which the NIH dominates basic research funding, but where big pharma firms also fund considerable research. The most important change in biotech research technology has been the US-sponsored Human Genome Project and associated new methods of genetic analysis and engineering that allow labs around the world to modify the biological underpinnings of living creatures to advance medicine and improve biological products and processes.

To measure collaboration patterns in the three fields we use publication data from the WoS. We identified all papers in the WoS database from 1990-2010 with at least one US

coauthor in journal subject categories *Physics, Particles & Fields; Nanoscience & Nanotechnology; and Biotechnology & Applied Microbiology*. From these papers, we identify teams by the names of coauthors and locate the authors by author affiliations. This sample includes 125,808 papers.

Using the location of the authors on each paper, we define four types of collaborations:

*US-Only Collaborations*, divided into *US collocated*, in which all US authors are in the same city; and *US Non-collocated* in which US coauthors are in at least two different cities.

*International Collaborations*, divided into *International/US collocated* in which US coauthors are in the same city with at least one foreign coauthor; and *International/US Non-collocated*, in which US coauthors are in two or more cities with at least one foreign coauthor.

Distinguishing between these forms of collaborations allows us to identify differences between papers with international collaborations and papers with collaborations in different locations, be they in the US or overseas, as well as between papers with collaborations across locations within the US. By focusing only on papers in which there is some US presence, our analysis may not generalize to papers written in which all authors are based outside the US; and by differentiating city location only for US coauthors, our findings do not address the potential effects of collocation or non-collocation of non-US based researchers on paper outcomes.

Figure 1 displays the proportion of papers in our four categories and the proportion with single authors in the three fields taken together in each year. The solid top line gives the share of papers in which a US-based author collaborates solely with authors collocated in the same city. It shows a marked decrease in collaborations between these persons from 1990 through 2000, which then stabilizes at about 40% of papers. The line labeled *Solo* shows the proportion of papers that are solo-authored. It drops from 20% to about 5% from 1990 to 2010. The line for International/US Collocated papers gives the share of papers for which at least one of the authors is in another country while all US authors are in the same city. It increases by 18 percentage points from 1990 to 2010. The line for International/US Non-collocated increases by about 5 percentage points from 1990 to 2010. Most of the increase in international collaborations was between US scientists based in one location and persons in another country. Overall, while papers with authors in different US cities increased less international collaborations, the data shows that increased geographic scope of collaborations involved more than crossing national boundaries.

To see whether the trend in collaborations varied noticeably among fields, Figure 2 displays the proportion of papers by collaboration type for the three fields separately. The data for particle physics in Figure 2a show the highest level of international collaborations, due presumably to the importance of particle accelerators and other equipment that are available at only some sites. Figure 2b and Figure 2c show that in nano and in biotech, the most common form of collaborations are US collocated teams, while international/US collocated collaborations are second most common and US non-collocated collaborations are third in frequency. International collaborations were roughly as common as US non-collocated collaborations in nano and biotech until the late 2000s, when international collaborations increased sharply. In all fields, the proportion of papers by sole researchers and by researchers collaborating in the same city falls.

The increase in international collaborations in our three fields resembles the patterns in National Science Board (2012) and in Adams (2013) for science more broadly. Similarly, the increased geographic dispersion of co-authorship in our fields reflects the pattern in science more broadly in the US as well.

## **2. Survey of Corresponding Authors**

To go beyond bibliometric data on collaborations, in August 2012 we conducted an on-line survey of the corresponding authors of papers published in 2004, 2007, and 2010 in the Web of Science Nano, Biotech, and Particle Physics subject categories with at least one US coauthor. We identified unique corresponding authors based on email addresses in these categories and selected one paper for each author, randomly choosing the paper from authors who had more than one paper in the database. Using the email address of the corresponding author, we sent a personalized email in English that invited them to complete the survey by clicking a link that connected them to the on-line survey instrument. If a paper had more than one corresponding author, we selected the one that appeared first. We sent 2 follow-up email reminders in August and September 2012. We used Qualtrics Survey Software and respondents accessed it from the Qualtrics server.

We customized each survey to ask the respondent about the specific collaboration and individual team members. The survey had 25 questions and was designed so that respondents could complete it in 10-15 minutes. The questions sought to discover how the team formed, how

it communicated and interacted during the collaboration, the contribution of each coauthor, types of research funding, and the advantages and disadvantages of working with the team. The survey also included an open-ended question for respondents to make comments. Several respondents sent emails with additional thoughts and information about the collaboration.

Between August 13, 2012 and August 20, 2012, we emailed a total of 19,836 individuals. Since some email addresses had expired, changed, or some individuals were deceased, the number of individuals who received the email is lower. We received 3,925 responses, which implies a response rate of 20% – a proportion that is in line with other surveys of scientists (Sauermann and Roach, 2013). For individuals who published their paper in the most recent year of our survey (2010), the response rate was 26%. Taking account of the proportion of emails that likely did not reach respondents, we estimate approximately 29% of recipients of emails answered them.<sup>3</sup>

The survey asked the respondent which country each coauthor was “*primarily based in during the research and writing*” of the article. This gives us a more accurate measure of whether teams are international than the WoS data, which are based on author affiliations at the time of publication, which can differ from those during the work either because affiliations change between the time of the research and the time of publication, or because some people have multi-country affiliations.

Table 1 compares the characteristics of collaborations in the papers we analyze to those in the full sample of WoS papers and those in the 2004, 2007, and 2010 WoS sample from which we drew the survey. Our final sample includes 3,452 respondents, which is lower in part than the returned responses due to the fact that some papers with US-addresses on the publication did not meet our requirement that at least one author be primarily based in the US at the time of the research. Our analysis uses the respondents’ information to define US collocated, US non-collocated and international teams.<sup>4</sup> The column giving the difference between the distribution of our sample in column 3 and the distribution of the WoS sample in column 2 shows that our

---

<sup>3</sup> Of those who received the email, 5,744 opened the survey, and 3,925 completed and submitted their answers. While we are unable to precisely count how many emails reached active mailboxes, based on the number of emails which “bounced” back from a sample of the messages sent, we estimate that approximately 32% of emails sent were undeliverable. Given this estimate, we approximate a response rate of 29% from the deliverable messages.

<sup>4</sup> Comparing the 34.01% in row “Int’l Collaboration Survey”, which is based on the respondent’s answers regarding the location of coauthors, and the 36.35% in “Int’l Collaboration” in Table 1 shows that using only reported author affiliations from publications overestimates the number of international teams by 2.35 percentage points.



survey sample is overrepresented by US collocated teams, the more recent publication year (2010), and publications from biotechnology.

### **3. Collaborations over Distance**

In what ways, if any, do papers with international collaborations differ from collaborations that occur solely in the US?

As others have found (e.g. Katz and Hicks, 1997; Rigby, 2009; Guerrero Bote, Olmeda-Gómez and de Moya-Anegón, 2013; Lancho Barrantes et al., 2012; and Adams, 2013), international collaborations tend to produce more highly cited papers than collaborations of persons in a single country. Taking all of our fields together gives a similar pattern, where the US is the single country to which we compare the international collaborations. We examined citations for papers published in 1990-2007 – dates chosen to allow time for papers to gain substantial numbers of citations. In this group US papers with foreign authors obtained 26.59 citations compared to 25.65 citations in which collaborations were solely with fellow residents of the US. Since US-authored papers average more citations than papers worldwide, it would have been reasonable to expect the opposite: fewer citations for US-based scientists collaborating with persons outside the country than for US-based scientists collaborating with other US scientists.

Does this mean that international collaborations per se produce better science as reflected in numbers of citations?<sup>5</sup>

We answer this question by comparing citations for papers with international collaborations and citations for papers with collaborations across locales in the US. If the observed international effect is due to something special about international collaborations, the average citations for international collaborations would exceed average citations for collaborations among non-collocated authors in the US as well as exceed the average citations for collocated authors in the US. Figure 3 shows the average number of citations for papers published between 1990 (with 21 years of potential citations) and 2007 (with three years of potential citations) for these three forms of collaboration. The number of citations varies over time, from approximately 30 for the older papers to 3-4 citations for the newer papers. In almost

---

<sup>5</sup> Citations measure the attention given to a paper, which is an imperfect measure of its scientific contribution since citation behavior can be driven by factors besides its contribution to knowledge (see e.g. Simkin and Roychowdhury, 2003). But it is still a sufficiently valuable indicator of the impact of a paper and is the most widely used measure in the science of science.

all years, papers with international collaborators and papers with non-located US collaborators have more citations than those published by collaborators in the same US city. But there is no clear pattern of differences in citations for papers coauthored by people in different US cities than for papers coauthored by people in the US and in a foreign location. Among papers published between 1998 and 2007, US non-located collaborations obtain more citations than international papers, but among papers published between 1990 and 1997, there is no clear difference. That cites per year between US non-located papers and international collaborations are reasonably similar and that both are notably larger than cites to US located papers suggests that the greater cites of international collaborations reflects multiple locations more than having authors across national borders.

We pursue the comparison of citations across types of collaborations for each of our three fields separately. The style of research in the fields differs greatly between particle physics, where empirical work often involves huge collaborations around particular pieces of equipment, and the smaller collaborations of nanotechnology and biotechnology research. This difference shows itself in the much higher average number of authors per paper in physics than in the other two fields (see Appendix Table A1). The difference is concentrated in the upper tail of the distribution of authors per paper. In particle physics, the upper 95<sup>th</sup> percentile of the number of authors per paper have 100 authors, while those in the 99<sup>th</sup> percentile have 523 authors – which far exceed the upper percentile numbers for authors in nano and biotech.

Reflecting the “big science” nature of some of the physics projects, the corresponding author on a physics paper with over 450 coauthors noted in our survey:

*This research was carried out as part of a very large collaboration in which every member gets authorship and this is listed in alphabetical order on our papers. The collaboration consists of scientists and engineers with a wide range of expertise - many primarily involved in designing, building and running instrumentation, and many analyzing data for various kinds of signal. This particular research was primarily carried out by myself, and the majority of the listed coauthors (including three of the selected authors in this survey) had no direct involvement in its preparation other than through collaboration membership.*

We next use regression analysis to examine the relation between the modes of collaboration and the number of researchers listed as authors and the number of references in a

paper, both of which tend to be positively related to citations. To the extent that the references influence the paper by providing information and ideas from other scientists they can be viewed as indicators of “invisible coauthors”, self-citations aside.

Table 2 records the regression coefficients and standard errors for regressions of numbers of coauthors and references on the type of collaboration and a year trend for each field. While there is a broad similarity in the estimated effect of the collaborations on the number of coauthors and references across the fields, there are also differences that presumably reflect differences in their research technologies. In all of the fields, the regression of number of coauthors on the dummy variable for whether or not the paper had an international co-author gives a positive coefficient on the dummy variable. But the magnitudes of the coefficients differ greatly. The estimated coefficient on international collaborations in particle physics in column 1 (43.8) shows that the number of authors on papers is much higher for those than for the US collaboration reference group; whereas the estimated coefficients for the relation between international collaborations and co-authors in nanotech and biotech are magnitudes smaller: 1.3 more authors on international papers than papers written by authors solely in the US for nano (column 2) and 2.2 more authors on international papers than US-only papers for biotech (column 3). The more detailed measures of collaborations in columns 4, 5, and 6 show that this difference is largely driven by international collaborations in which the US scientists doing particle physics are from many locations as well. This reflects the big science nature of empirical particle physics, where huge numbers of collaborators work together with massive instruments and machines compared to the smaller lab science of nano and biotech.

The regressions in columns 7-9 show greater differences among the fields in the number of references on international papers relative to US non-located papers, and differences among the fields in the relation between numbers of co-authors and numbers of references. In particle physics, numbers of references for papers with international collaborations exceed those for US non-located papers, which in turn exceed both those for US located international papers and those US located (column 7). In biotech, numbers of references for papers with international collaborations and non-located US collaborations exceed those for US non-located papers, which in turn exceed those for US located papers (column 9). A potential explanation is that persons in a given location are more likely to cite papers written in their location, so that the greater the number of locations, the greater the number of references. But

the regression for number of references in the nano papers shows a different relation between references and collaborations (column 8). Finally, the estimated coefficients on the number of authors also shows no consistent pattern among the fields: negligible effects for particle physics (potentially because the number of authors can be extremely high), slight negative effects in nano, but substantial positive effects in biotech (column 9).

All told, Table 2 shows that simple comparisons of papers with international and national collaborations can present a misleading picture about the science involved in various types of collaborations. The collaborations can involve huge differences in the numbers of coauthors and differing relations to the numbers of references. Given these results, we examine the relation between the citations to a paper and the form of collaboration separately for each field using a regression that includes the number of coauthors and the number of references in the paper. To deal with the life cycle of citations in which the number of citations increases sharply in the first 5 to 7 years after publication and then grows more slowly we include dummy variables for the year the paper was published as well.

Tables 3a – 3c gives the results of this analysis. Column 1 of each of the tables estimates the difference in citations between international papers and US only collaborations. The estimates show a disparate pattern across the fields: an insignificant positive relation between international collaborations and citations for particle physics; a negative relation in nano; and a positive relation in biotech. Column 2 of each table adds the number of coauthors to the regression. In each of the fields, the addition of numbers of authors reduces the coefficient on international collaborations. In biotech it turns the coefficient from positive to negative.<sup>6</sup> With the addition of numbers of references in column 3, the estimated relation of international collaborations to citations is significantly negative in all three fields. The disaggregation of types of collaborations in columns 4 in tables 3a-3c shows sufficiently weak and different patterns across the fields to suggest that there is nothing universal in the link between international collaborations and ensuing citations to papers.

All told, the regression analysis in Tables 2 and 3 document the changing patterns of cooperation across locations in the three fields and their disparate relation with citations. While

---

<sup>6</sup> To see if this is a more general pattern, we ran similar regressions for other scientific fields in the WoS and find variation across fields in the difference between citation rates for international collaborations and domestic collaborations; the addition of the number of coauthors to citation regressions reduces the coefficient on international collaborations in almost all fields.

invaluable as descriptions about collaborations, such bibliometric analysis cannot, however, provide insight into the ways collaborating scientists work together to conduct the research that leads to published papers. To gain insight into what goes on in collaborations, we turn to the survey of corresponding authors described in Section 1.

#### 4. Survey Evidence

*“I think the best example of collaboration I have done is...where all the authors are from different countries and we met at the Bellagio Conference Center of the Rockefeller Foundation.”*

*“I think that it is absolutely indispensable to meet people in person to have effective collaborations.”*

*“Skype was not available...at the time we completed this work. We now use Skype or ITV connection to meet and discuss data with collaborators on a weekly basis.”*

*“The international collaboration worked so well because of my frequent trips to Brazil during the project.”<sup>7</sup>*

For scientists to collaborate, they must meet and decide to work together, communicate during the collaboration, and combine their knowledge and skills to create sufficient new knowledge to generate a publishable paper.

#### Meeting and Communicating

We asked corresponding authors to answer the following question about their coauthors: *“How did you FIRST come in contact with each of these coauthors?”* For papers with up to six authors, we asked about each coauthor. For papers with more than six we asked about the first and the last authors if they were not the corresponding author and about randomly selected authors from the list of coauthors to obtain information on a maximum of six collaborators.

Figure 4 displays the proportion of persons of each collaboration type who the corresponding author first met as advisor-student/post-doc; colleagues in the same department/institution; through contact without an introduction; at a conference, seminar or other meeting; or by visiting the department/institution. The figure shows that for all forms of

---

<sup>7</sup> The four quotes are based on comments from the open-ended section of our survey.

collaboration, most first meetings occurred when the corresponding author and the other person worked in the same institution. For papers written in the same location, the predominant contact was through advisor-student or post-doc relationships but that over one third of the meetings came about as colleagues. For papers with authors from other US locations or foreign locations, the corresponding author met them through working in the same place, primarily as a colleague, but with nearly ten to sixteen percent meeting the person as a visitor. Conferences also accounted for a substantial proportion of the first meetings between corresponding authors on papers written with persons in other US locations or in foreign locations.<sup>8</sup> Overall, Figure 4 shows broad similarity in the mode of meeting between non-located US authors and in the mode of meeting between US and foreign-located authors compared to the mode of meeting for coauthors in US located collaborations.

We asked corresponding authors the frequency with which they communicated with one or more of their coauthors from “every week” to “never”. Because collaborations that include persons in the same locale and persons in other locales as the corresponding author allow the corresponding author to meet face-to-face easily with some coauthors but only infrequently with coauthors in other locations, the question does not pin down differences associated with distance. To overcome this problem, we show in Figure 5 modes of communication between coauthors on two-authored papers, which differentiate properly communication between located, non-located US, and foreign coauthors.

The results show that the corresponding author relies extensively on face-to-face meetings when all authors are in the same location. But Figure 5 also shows that while face-to-face meetings are much lower for authors across distances, such meetings are still frequent. Among the 2-author papers, just over 50% of corresponding authors on international teams report meeting face-to-face at least a few times per year, while 64% of those on US non-located papers reported face-to-face meetings at least a few times a year. By contrast, the figure shows no noticeable differences in using e-mail by distance. Corresponding authors in all forms of collaborations use e-mail frequently to communicate with their collaborators, approximately 40 weeks during the year. There are substantial differences in use of telephone

---

<sup>8</sup> The time series data in Appendix Figure A1 and A2 show that conferences have become a less important way to meet future coauthors, while students/postdocs have become more important, possibly due to their increased importance in the scientific production process.

vs. Internet (e.g. Skype) between US-based teams and international teams that are readily explained by the differential in cost of international and within US telephone calls.

Our survey findings that face-to-face meetings are important in both the initiation of research collaborations and the working of distant collaborations are consistent with evidence on the role of co-location in the formation of research collaborations (such as Boudreau et al. 2014) and the need for periodic co-location to maintain the effectiveness of distant collaborations even with advances in long-distance communication technologies (see e.g. Olson and Olson, 2000; Olson and Olson, 2003; Cummings and Kiesler 2005).

### **What Coauthors Bring to Collaboration**

To understand what factors helped produce the collaborations, we asked the corresponding author to specify the unique contribution of each team member. Our question was *“Did any of the team members working on this article (including yourself) have access to one of the following resources that the other team members did NOT have which made it important for you to all work together on this topic?”* The possible choices were: access to data, material or components; data, material or components protected by intellectual property; a critical instrument, facility or infrastructure; funding; or unique knowledge, expertise or capabilities.

Figure 6 shows that the major factor cited for all collaborations was “unique knowledge, expertise, capabilities”. That access to specialized human capital seems to drive collaborations, whether US or international, implies that a theory of collaboration should focus on the complementarity of skills and knowledge of collaborators just as the theory of trade focuses on comparative advantage in creating trade among countries. But there are differences in the importance of other factors across forms of collaboration. Non-located and international teams were more likely to have a coauthor contributing data, material or components than US located teams— a pattern that has increased over time (see Appendix Figure A3).

While most corresponding authors reported the contribution and role of their coauthors, those on huge collaborations told a different story. As one respondent remarked, “Many of the questions are hard to translate to the field of experimental particle physics, where an international collaboration of hundreds of scientists work on the same project with funding from many countries. I can only guess, where the funding from each of the ~300 coauthors comes from, many of whom I have not even met. The published research is primarily the work of a

single person (myself), but would not have been possible without having access to custom software and data provided by the collaboration."

Finally, taking advantage of the unique identification of authors in two-authored papers, we compare the specific contributions of foreign-located coauthors and domestic coauthors on those papers. US and foreign coauthors were equally likely to contribute "unique knowledge, expertise, or capabilities" and "data, material or components protected by intellectual property". Foreign coauthors are slightly more likely to contribute access to "data, material or components" or "a critical instrument, facility or infrastructure" while the US coauthor was slightly more likely to contribute funding.

### **Advantages and Challenges**

To assess the effects of the different forms of collaboration on the production and output of scientific activity, we use our survey, where we asked the corresponding authors their views of the advantages and challenges on their collaboration, and the bibliometric data, where we estimated a regression model linking the number of citations to a paper to the attributes of the collaboration reported on the survey.

Table 4 summarizes the responses of corresponding authors on the advantages and challenges of the collaborations. It records the average score on a five-point scale of agreement (5) or disagreement (1) with statements regarding the attributes of the collaboration. The corresponding authors agreed that their collaboration had substantial advantages in harnessing human capital to produce a scientific outcome. "Complementing our knowledge, expertise and capabilities" and "learning from each other" are the only items with average scores greater than 4 in the table. The next highest score was that collaborations made the research experience more pleasant. There is little variation here in the responses between US non-located and international teams. Corresponding authors on the both of those collaborations gave modestly higher scores to the knowledge advantages than the located teams. Similarly, all three groups ranked highly "Gaining access to data, material or components", with the highest assessment coming from the corresponding authors of US non-located teams.

The corresponding authors of international teams gave higher scores to the advantage of "Our research reached a wider audience" than did the corresponding authors of US-non-located teams who in turn gave higher scores than the corresponding authors of US located



only teams. Viewing “wider audience” in terms of the geographic distribution of citations, this suggests that the wider geographic distribution of authors, the wider is the distribution of citations, possibly even among papers with the same numbers of citations.

Regarding the challenges of collaborations, US non-located and international teams reported similarly that there was “Insufficient time for communication”, “Problems coordinating with team members’ schedules”, and “Insufficient time to use a critical instrument, facility or infrastructure” than did US co-located teams. As in the bibliometric analysis in section three, geographic location appears to be more than national boundaries in the way teams operated.

We also asked whether the corresponding authors viewed teams as having the optimal size. The responses, given in Appendix Table A2, show that most corresponding authors viewed their team as having the right size. Presumably the principal investigator(s) would have modified the team if they did not think that was the case. But there are some differences by collaboration type. US located teams were more likely to say that they needed additional collaborator (7.58% vs. 3.48% and 3.38% for US non-located and international); whereas international teams were more likely to say that fewer team members were needed (6.67% vs. 3.37% for US located). Reflecting the role of government policies, twenty four percent of the international teams received funding aimed at supporting cross-country collaboration, with 6.65% receiving US government funding, 4.64% receiving EU funding, and the remainder from other government sources.

As our second way to assess how the attributes of collaborations affect outcomes we added the corresponding authors' descriptions of the collaboration to the Table 3 regressions of the number of citations on attributes of papers. Because publication of the paper preceded the survey, some of the corresponding author views of the collaboration will presumably have been affected by the success of the paper, which would give a distorted view of the link from collaboration to outcome. To deal with this problem, we limit analysis to the survey responses that seem least prone to be affected by the outcome – relatively objective questions about the way corresponding authors met coauthors, what coauthors contributed, and funding support.

Table 5 gives the results of this analysis. Columns 1 and 2 replicate the regression estimates in Table 3 for the dichotomous international collaboration variable. The results in Table 5 show some differences in the regression coefficients from that found in the larger WoS sample. The positive coefficient on international collaborations in column 1 in Table 5 is larger

than the coefficient in the comparable regression using the larger WoS sample papers in our 3 fields. The coefficients on the number of coauthors and number of references variables are positive and significant in column 2 of Table 5 but the coefficient on coauthors is larger than that of references, contrary to the result in the larger WoS sample. Subject to these differences, which suggest some modest differences between the papers of respondents to the survey and the population of papers, the estimated coefficients on the survey variables in columns 3, 4 and 5 tell a clear story. They show that papers in which at least one coauthor met at a conference had higher citations; that papers for which a coauthor contributed funding had lower citations, and that papers that got funding specifically for cross-country collaborations had lower citations.<sup>9</sup> The natural interpretation of these patterns is that collaborations based on ideas or relations developed at conferences produce more cited and potentially better science than collaborations based on funding.

## 5. Toward an Economics of Scientific Collaborations

Scientific collaborations have become increasingly important in scientific research, but the nature of collaborations, their determinants, effects on scientific outcomes, and the incentives that drive scientists to collaborate or not, and with whom to collaborate, is not well understood. From the perspective of economics, collaborations occur because they enhance scientific productivity, but collaborations have costs as well as benefits – the costs of communicating ideas and coordinating disagreements among collaborators and the expenses of getting them together in one place or linking them with unique data and equipment. But arguably the biggest problem collaborations must solve in order to succeed is to find ways to divide the credit for enhanced productivity among persons, so that each collaborator prefers working in the team rather than by themselves, where they gain full credit for research outcomes.

This study has linked a unique survey of corresponding authors in particle physics, nanotechnology, and biotechnology with bibliometric data from the same fields to identify some of the key empirical relations in the growth of collaborations from the 1990s to 2010. The bibliometric data shows that the share of papers with a US address written by collaborators in different locations increased in the period studied and that the largest increase occurred across

---

<sup>9</sup> We also estimated the model including dummies for whether the corresponding author didn't view the team size as optimal, and an average of the scores assessing the advantages and disadvantages to the collaboration, but found no effect of these measures on citations.

country lines, followed by collaborations in different locations in the US. Commensurately, the share of papers written by single scientists or by groups of scientists in single locations declined.

The survey of corresponding authors shows similarities in the way collaborators first meet and later communicate and work together among different types of collaborations, save for very large physics projects. The bibliometric data shows that the number of collaborators on a paper is positively associated with the numbers of citations to the paper, but the data is mixed on whether international collaborations are more productive in terms of citations than domestic collaborations. In biotech, international collaborations obtain more citations than domestic collaborations; in nanotech they obtain fewer citations; while particle physics shows no significant differences between international and domestic collaborations. In all three fields, papers with the same number of coauthors had lower citations if they were international collaborations, which suggests that a main advantage of international collaborations for US authors is that they allow researchers to increase the number of collaborators more easily than if the supply of potential coauthors was limited to US-based scientists.

The data are thus consistent with the notion that collaborative work has greater productivity or impact as reflected in citations, but do not give a clear message about whether the fastest growing form of collaborations – that across country lines – has any productivity edge over collaborations across space within the US. The most likely reason for the rapid growth of international collaborations is the more rapid growth of science and engineering PhDs and researchers in other countries than in the US, which creates a large supply of potential collaborators overseas.

Wherein does the productivity advantage of collaborations lie? Viewing science as an aggregate process for producing new knowledge, the most plausible answer is that the knowledge base has become increasingly complex and specialized (Jones, 2011), and thus scientific advances require increased numbers of researchers combining their skills and expertise. Consistent with this, our survey of corresponding authors shows that access to specialized human capital is the main driver of collaborations. The growing number of references within papers suggests that each forward step in science builds on a large base of previous knowledge. And the positive link between numbers of references and citations suggests that the greater the knowledge that goes into a paper, the greater the scientific contribution of the paper – at least to the extent that these measures are a valid “paper trail” of flows of knowledge. All of which is

consistent with the view that the productivity advantage from collaborations depends on the combination of ideas/knowledge from persons with different expertise,<sup>10</sup> though it does not prove the validity of this interpretation.

But, as noted, collaborations have costs as well as benefits, and decisions to collaborate involve balancing the benefits against the costs. On one side are problems of coordinating the ideas of persons with different expertise or viewpoints or who are in different locations, and the expenses and difficulty of getting collaborators together or linking them with data and key pieces of machinery. Our survey finding that researchers meet most collaborators through personal connections made at their institution or, to a lesser extent, at conferences, suggests that there is some role of chance in creating collaborations. Most important, the survey finding that corresponding authors view face-to-face meetings as critical in their collaborations, and hold them relatively often, suggests that the improvements in communication and reductions in their cost do not fully substitute for human interactions in collaborations.

To an individual researcher, the biggest issue in collaboration is getting credit for a joint production. In a one-author paper, the one takes credit or blame. In a two-author paper, many fields adhere to the convention that the senior person's name comes last and the junior person comes first, which potentially gives substantial credit to each. Freeman and Huang (2014) find that the impact factor of the placement of the paper and citations received depend more on the characteristics of the last named, typically senior, author rather than of the first named author. The senior person thus appears to play a greater role in gaining attention to the research, which can help the junior author as long as other scientists view the first author as more than a pair of hands in the lab. In papers with more than two authors, the decision on who is the first author and the placement of the non-first or non-last authors can create disputes. On papers having huge numbers of names where tasks are highly specialized, the credit to a given author is presumably related to their specialty, much like the credits that give the names of specialists in a movie production. Only people who know how the research (movie production) proceeded and what the particular person's function was would understand how to evaluate their contribution in the author (credit) list.

---

<sup>10</sup> Weitzman's (1998) model of the growth of useful knowledge from combining the growing supply of past ideas and knowledge in new ways offers a way to structure such a model.

From the perspective of economic rationality, the decision of scientists to collaborate depends on both the productivity of the collaboration and the distribution of credit. To get some notion of the interplay of the factors, consider the situation in which a scientist compares the value of collaborating on a paper with one or more other scientists to writing a paper by themselves. On the productivity/citation side, assume that a paper with  $N$  collaborators gains proportionately more citations ( $C$ ) than a solo-authored paper according to a linear productivity parameter  $p > 1$  that links citations to numbers of authors by  $C = pN$ . If a single authored paper gets  $p$  citations, then this function gives two authored papers  $2p$  citations, a paper with three authors  $3p$  citations, and so on.

But whereas each author gets full credit for an individual paper, they get only partial credit for joint work. Assume that the science community allocates credit for joint work with a citation crediting function  $\gamma(n)$  that has the value 1 for  $\gamma(1)$  and in which the derivative  $\gamma'(n) < 0$  so that  $\gamma(1)=1 > \gamma(2) > \gamma(3)$  and so on. The only restriction on the citation crediting function is that each author gets less credit the larger the collaboration. Someone seeking to maximize the number of citations credited to them would collaborate only if  $p\gamma \geq 1$  – that is, if the gain in productivity from the collaboration exceeds the loss of credit associated with  $\gamma$ . If the crediting function was based on simple fractionalization of credit, so that each author in a two authored paper would be credited with  $1/2$  of the paper and thus  $1/2$  of the cites,  $p$  would have to exceed 2 for the two-authored paper to be worthwhile. Similarly,  $p$  would have to exceed  $n$  for an  $n$ -sized collaboration to be attractive. But estimates of the extent to which citations increase with number of authors falls far short of such proportionality. Depending on field and specification, our Table 3 estimates showed that additional authors raises citations by at most 1-2 citations per additional author. In this case, if a solo-authored paper gained 10 citations, here would be little incentive to write a joint paper that gained 12 citations for which each author obtained credit for just 6 citations.

Why then have collaborations increased so much in the sciences?

One possibility is that scientists who collaborate with others are able to write so many more papers through division of labor than by themselves that the increased number of papers offsets the lower credit set by the crediting function. For most scientists, this seems unlikely. The average number of collaborators in science articles has roughly doubled in the past 4-5 decades, while the number of papers written per researcher has not shown any such doubling.

Most of the increase in papers over time has been associated with increased numbers of researchers rather than increased papers per researcher.

The other possibility, which we view as the likely solution to the question, is that the crediting function diverges greatly from fractionalization. Fractionalization imposes the constraint that the sum of credits to all authors is one, but there is no law that rules out allocating credit differently. The first author of a highly successful paper gains lots of credit for their contribution. The last author can also gain lots of credit for the different contribution they made. The authors in the middle of the author list presumably gain some credit, but less than the other two authors. The purpose of the crediting is to propel the careers of persons who are part of a successful team activity. Thus, we would expect first and last authors to benefit most from a successful collaboration in their future careers, and for intermediate authors to benefit proportionate to their role on the research, subject to the imperfections in markets and market information. While we cannot test this interpretation with our data, it is testable with information on the future careers of persons who work on papers with different numbers of collaborators. Our prediction would be that two coauthors of a well-cited paper would gain more in their careers from the joint paper than if each had written a solo paper with half as many citations.

Finally, to the extent that the interplay between the productivity of working with other scientists and the distribution of credit affect collaboration decisions as hypothesized above, we would expect to find at most modest differences between the nature and effects of collaborations across national borders as within the US, as our survey and WoS data analysis seem to show.

## References

- Adams, J. (2013). "Collaborations: The Fourth Age of Research." *Nature*, 497(7451), 557-560.
- Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). "Scientific Teams And Institutional Collaborations: Evidence From US Universities, 1981–1999." *Research Policy*, 34(3), 259-285.
- Agrawal, A., A. Goldfarb. (2008). "Restructuring Research: Communication Costs and the Democratization of University Innovation." *American Economic Review* 98(4):1578–1590.
- Barrantes Bárbara S. Lancho, Vicente P. Guerrero Bote, Zaida Chinchilla Rodríguez, Félix de Moya Anegón (2012). "Citation Flows in the Zones of Influence of Scientific Collaborations" *Journal of the American Society for Information Science and Technology*, Volume 63, Issue 3, pages 481–489, March 2012.
- Boudreau, K., Brady, T., Ganguli, I., Gaule, P., Guinan, E. C., Hollenberg, H. & Lakhani, K. R. (2014). "A Field Experiment on Search Costs and the Formation of Scientific Collaborations," (August 24). available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2486068](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2486068)
- Catalini, C., Fons-Rosen, C., Gaule, P. (2014). "Air Travel Costs and Scientific Collaborations." Mimeo. MIT Sloan School of Management, unpublished manuscript.
- Cummings, J. N., & Kiesler, S. (2005). "Collaborative Research Across Disciplinary and Organizational Boundaries." *Social Studies of Science*, 35(5), 703-722.
- deB. Beaver, Donald (2004). "Does Collaborative Research Have Greater Epistemic Authority?" *Scientometrics*, 60(3), 399-408.
- Freeman, Richard B. (2010). "Globalization of Scientific And Engineering Talent: International Mobility of Students, Workers, and Ideas and The World Economy." *Economics Of Innovation And New Technology*, Volume 19, issue 5, 201 pp. 393-406.
- Freeman, R.B. and Huang, W. (2014). "Collaborating With People Like Me: Ethnic Co-authorship within the US." NBER WP 19905
- Guerrero Bote, Vicente P. Carlos Olmeda-Gómez Félix de Moya-Anegón (2013). "Quantifying the Benefits of International Scientific Collaboration," *Journal of the American Society for Information Science and Technology*, Volume 64, Issue 2, pages 392–404, February 2013.
- Hsu, J.W., & Huang, D.W. (2011). "Correlation Between Impact and Collaboration." *Scientometrics*, 86(2), 317–324.
- Jones, B. F., Wuchty, S., & Uzzi, B. (2008). "Multi-University Research Teams: Shifting Impact, Geography, And Stratification In Science." *Science*, 322(5905), 1259-1262.

Jones, Benjamin F. (2011). "As science evolves, how can science policy?" *Innovation Policy and the Economy*, Volume 11. University of Chicago Press, 2011. 103-131.

Katz, J.S. And D. Hicks (1997). "How Much Is a Collaboration Worth? A Calibrated Bibliometric Model," *Scientometrics*, 40:3, 541-554.

Lancho Barrantes, Bárbara S., Vicente P. Guerrero Bote, Zaida Chinchilla Rodríguez, and Félix de Moya Anegón (2012) "Citation Flows in the Zones of Influence of Scientific Collaborations," *Journal of the American Society for Information Science and Technology* 63(3): 4819.

Lawani, S. M. (1986). "Some Bibliometric Correlates Of Quality In Scientific Research," *Scientometrics*, 9:1-1 13-25.

Lee K, Brownstein JS, Mills RG, Kohane IS (2010). "Does Collocation Inform the Impact of Collaboration?" *PLoS ONE* 5(12): e14279. doi:10.1371/journal.pone.0014279

National Science Board (2012). *Science and Engineering Indicators*, 2012.

Olson, G., & Olson, J. (2003). "Mitigating the Effects of Distance on Collaborative Intellectual Work." *Economics of Innovation and New Technology*, 12(1), 27-42.

Olson, G. M., & Olson, J. S. (2000). "Distance Matters". *Human-Computer Interaction*, 15(2), 139-178.

Rigby, John (2009). "Comparing the Scientific Quality Achieved By Funding Instruments For Single Grant Holders And For Collaborative Networks Within A Research System: Some Observations," *Scientometrics* 78:1, 145-164.

Sauermann, H., & Roach, M. (2013). "Increasing Web Survey Response Rates In Innovation Research: An Experimental Study Of Static And Dynamic Contact Design Features." *Research Policy*, 42(1), 273-286.

Scellato, G., Franzoni, C., & Stephan, P. (2012). "Mobile Scientists and International Networks," NBER Working Paper 18613.

Simkin, M. V., & Roychowdhury, V. P. (2003). "Read Before You Cite!" *Complex Systems*, 14, 269-274.

Thomson Reuters (2012). *Web of Science*®, prepared by THOMSON REUTERS®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA: © Copyright THOMSON REUTERS® 2012. All rights reserved.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). "The Increasing Dominance Of Teams In Production Of Knowledge." *Science*, 316(5827), 1036-1039.





## TABLES

Table 1. Distribution of papers by characteristics, Web of Science Papers and Survey Respondents

	(1) Papers, 1990-2010	(2) Papers in 2004, 2007, 2010	(3) Survey Sample, Papers in 2004, 2007, 2010	(3)-(2) Difference
<i>Collaboration Type</i>				
US Collaboration Only	66.29	63.65	62.25	-1.4
US Collocated	44.81	41.56	46.84	5.28
US Non-Collocated	21.47	22.09	15.41	-6.68
Int'l Collaboration	33.71	36.35	37.75	1.4
Int'l/US Collocated	24.04	26.04	26.94	0.9
Int'l/US Non-Coll.	9.68	10.31	10.81	0.5
Int'l Collaboration Survey			34.01	
<i>Year</i>				
2004	6.08	25.38	18.42	-6.96
2007	8.05	33.61	29.46	-4.15
2010	9.83	41.01	52.11	11.1
<i>Field</i>				
Particle Physics	25.19	21.75	19.55	-2.2
Nano	23.82	32.85	30.5	-2.35
Biotechnology	50.99	45.40	49.94	4.54
N	125,808	30,141	3,452	

Notes: (1) includes all papers in the Web of Science with more than 1 author, at least one US coauthor, and with journal subject categories of Physics, Particles & Fields; Nanoscience & Nanotechnology; and Biotechnology & Applied Microbiology, published from 1990-2010. (2) includes those papers in 2004, 2007, and 2010. (3) includes the respondents to our survey, which was a sample based on unique corresponding authors appearing in (2) that had more than 1 author.

Table 2: Estimated Relation Between Number of Coauthors and Number of References on Papers by Nature of Collaboration, By Field

	<u>Coauthors</u>			<u>Coauthors</u>			<u>References</u>		
	Particle Physics (1)	Nano (2)	Biotech (3)	Particle Physics (4)	Nano (5)	Biotech (6)	Particle Physics (7)	Nano (8)	Biotech (9)
US Collaboration Only									
US Collocated									
US Non-Collocated				2.654** (0.150)	1.450** (0.033)	1.688** (0.029)	3.453** (0.377)	-0.879** (0.232)	0.727** (0.179)
Int'l Collaboration	43.776** (0.924)	1.331** (0.032)	2.168** (0.040)						
Int'l/US Collocated				12.017** (0.641)	1.458** (0.033)	1.973** (0.032)	4.737** (0.313)	-0.963** (0.272)	0.275 (0.189)
Int'l/US Non-Collocated				99.983** (2.091)	3.075** (0.073)	5.015** (0.126)	4.590** (0.400)	0.168 (0.400)	3.131** (0.359)
No. Coauthors							0.001 (0.001)	-0.060 (0.042)	0.435** (0.031)
Year Trend	-0.214* (0.094)	0.039** (0.003)	0.078** (0.002)	-0.183* (0.090)	0.038** (0.003)	0.064** (0.002)	0.796** (0.024)	1.491** (0.024)	0.535** (0.013)
Constant	433.018* (188.702)	-73.670** (6.828)	-151.290** (4.713)	368.918* (179.207)	-71.578** (6.581)	-125.213** (4.484)	-1.6e+03** (47.520)	-3.0e+03** (48.749)	-1.0e+03** (26.100)
R2	0.055	0.068	0.091	0.170	0.144	0.159	0.046	0.116	0.044
Nb. of Obs.	31,690	30,761	64,153	31,690	30,761	64,153	31,690	30,761	64,153

Notes: + p < 0.10, \* p < 0.05, \*\* p < 0.01, OLS estimation. Includes all papers in the Web of Science with more than 1 author, at least one US coauthor, and with journal subject categories of Physics, Particles & Fields; Nanoscience & Nanotechnology; and Biotechnology & Applied Microbiology, published from 1990-2010.

Table 3a: The Estimated Relation Between Number of Citations to a Paper and the Type of Collaboration That Produced the Paper, Particle Physics

	(1)	(2)	(3)	(4)
US Collaboration Only				
US Collocated				
US Non-Collocated				1.664 <sup>*</sup> (0.691)
Int'l Collaboration	0.718 (0.469)	0.096 (0.452)	-1.212 <sup>**</sup> (0.464)	
Int'l/US Collocated				-1.418 <sup>**</sup> (0.532)
Int'l/US Non-Collocated				1.402 (0.856)
No. Coauthors		0.014 <sup>**</sup> (0.003)	0.014 <sup>**</sup> (0.002)	0.010 <sup>**</sup> (0.002)
No. References			0.398 <sup>**</sup> (0.017)	0.396 <sup>**</sup> (0.017)
Constant	24.030 <sup>**</sup> (1.953)	24.031 <sup>**</sup> (1.945)	15.404 <sup>**</sup> (1.894)	14.817 <sup>**</sup> (1.901)
Year FE	Yes	Yes	Yes	Yes
R2	0.030	0.031	0.072	0.073
Nb. of Obs.	31,690	31,690	31,690	31,690

Notes: +  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , OLS estimation. Sample is all papers in the Web of Science with more than 1 author, at least one US coauthor, and with a journal subject category of Physics, Particles & Fields, published from 1990-2010.

Table 3b: The Estimated Relation Between Number of Citations to a Paper and the Type of Collaboration That Produced the Paper, Nanotechnology

	(1)	(2)	(3)	(4)
US Collaboration Only				
US Collocated				
US Non-Collocated				-3.971** (0.423)
Int'l Collaboration	-2.300** (0.358)	-3.732** (0.388)	-3.637** (0.387)	
Int'l/US Collocated				-4.849** (0.470)
Int'l/US Non-Collocated				-6.305** (0.621)
No. Coauthors		1.074** (0.083)	1.110** (0.080)	1.294** (0.085)
No. References			0.295** (0.068)	0.293** (0.068)
Constant	26.252** (4.749)	21.712** (4.683)	14.660** (4.805)	14.747** (4.819)
Year FE	Yes	Yes	Yes	Yes
R2	0.039	0.045	0.068	0.070
Nb. of Obs.	30,761	30,761	30,761	30,761

Notes: +  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , OLS estimation. Sample is all papers in the Web of Science with more than 1 author, at least one US coauthor, and with a journal subject category of Nanoscience & Nanotechnology, published from 1990-2010.

Table 3c: The Estimated Relation Between Number of Citations to a Paper and the Type of Collaboration That Produced the Paper, Biotechnology

	(1)	(2)	(3)	(4)
US Collaboration Only				
US Collocated				
US Non-Collocated				1.109 <sup>*</sup> (0.531)
Int'l Collaboration	1.800 <sup>**</sup> (0.597)	-1.466 <sup>*</sup> (0.680)	-1.583 <sup>*</sup> (0.677)	
Int'l/US Collocated				-2.138 <sup>**</sup> (0.647)
Int'l/US Non-Collocated				2.394 (1.891)
No. Coauthors		1.506 <sup>**</sup> (0.103)	1.412 <sup>**</sup> (0.101)	1.333 <sup>**</sup> (0.110)
No. References			0.193 <sup>**</sup> (0.015)	0.191 <sup>**</sup> (0.015)
Constant	34.629 <sup>**</sup> (2.047)	29.522 <sup>**</sup> (2.054)	24.805 <sup>**</sup> (2.075)	24.917 <sup>**</sup> (2.067)
Year FE	Yes	Yes	Yes	Yes
R2	0.025	0.032	0.036	0.036
Nb. of Obs.	64,153	64,153	64,153	64,153

Notes: +  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , OLS estimation. Sample is all papers in the Web of Science with more than 1 author, at least one US coauthor, and with a journal subject category of Biotechnology & Applied Microbiology, published from 1990-2010.

Table 4. Advantages and Challenges to Working with the Team

	US Collocated	US Non- Collocated	Int'l
<u>Advantages</u>			
Learning from each other	4.26	4.33	4.36
Complementing our knowledge, expertise and capabilities	4.39	4.58	4.57
Gaining access to data, materials or components	3.21	3.56	3.32
Gaining access to data, materials or components protected by IP	2.14	2.30	2.29
Our research reached a wider audience	3.24	3.37	3.48
The research experience was more pleasant	3.96	3.92	4.02
<u>Challenges</u>			
Insufficient time for communication	1.82	2.13	2.11
Less flexibility in how the research was carried out	1.73	1.99	1.93
Unable to unequivocally portray my contribution	1.55	1.59	1.65
Problems coordinating with team members' schedules	1.96	2.18	2.11
Insufficient time to use a critical instrument, facility or infrastructure	1.45	1.67	1.67
Observations	1,693	585	1,174

Notes: Respondents were asked to indicate their level of agreement with these statements regarding the main advantages/disadvantages of “*carrying out the research for this article with your team members*”, where 5 = Agree, 1 = Disagree.

Table 5: The Estimated Relation Between Number of Citations to a Paper and the Type and Characteristics of Collaboration, Survey Sample

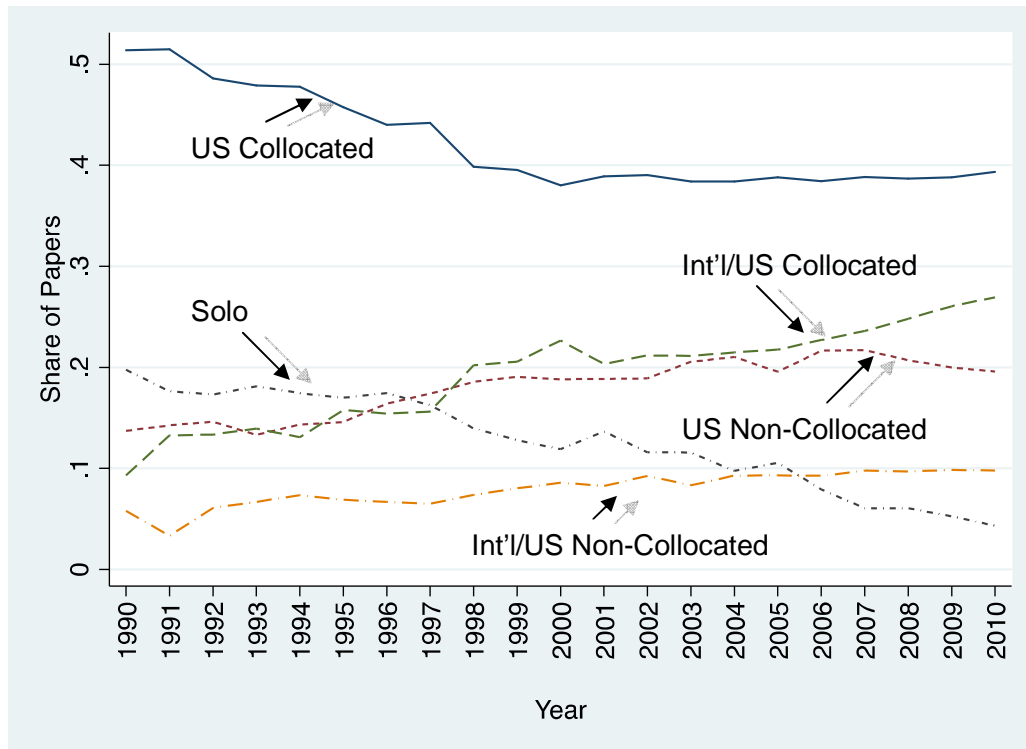
	(1)	(2)	(3)	(4)	(5)
US Collaboration Only					
US Collocated					
US NonCollocated			-0.355 (0.779)	0.434 (0.779)	0.444 (0.773)
Int'l Collaboration	0.878 <sup>+</sup> (0.529)	0.192 (0.538)	-0.579 (0.649)	0.370 (0.600)	0.495 (0.586)
No. Coauthors		0.161* (0.066)	0.157* (0.066)	0.160* (0.066)	0.161* (0.066)
No. References		0.099** (0.015)	0.098** (0.015)	0.099** (0.015)	0.098** (0.015)
<i>How They Met</i>					
Advisor-Stu./Postdoc			-0.734 (0.656)		
Colleagues			0.592 (0.547)		
Visiting			0.703 (0.877)		
Conference			2.939** (0.993)		
No introduction			0.575 (0.890)		
<i>Coauthor Contributions</i>					
Knowledge, etc.				0.498 (0.682)	
Funding				-1.327* (0.553)	
Data, etc.				-0.305 (0.520)	
IP Data, etc.				0.124 (0.630)	
Instrument, etc.				0.166 (0.567)	
Cross-country funding					-1.207* (0.610)
Constant	17.433** (2.497)	14.655** (2.513)	14.654** (2.593)	14.925** (2.607)	14.676** (2.548)
R2	0.076	0.114	0.119	0.116	0.115
Nb. of Obs.	3,452	3,452	3,452	3,452	3,452

Notes: +  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , OLS estimation. All regressions include year, field, and year x field fixed effects. Sample is the survey sample described in Section 2. “How They Met” and “Coauthor Contribution” variables are dummies indicating whether any coauthor on the team met that way/contributed the resource.



## FIGURES

Figure 1. Share of Papers by Collaboration Type



Notes: Includes all papers in the Web of Science with at least one US author, and with journal subject categories of Physics, Particles & Fields; Nanoscience & Nanotechnology; and Biotechnology & Applied Microbiology, published from 1990-2010.

Figure 2a: Share of Papers by Collaboration Type, Particle Physics

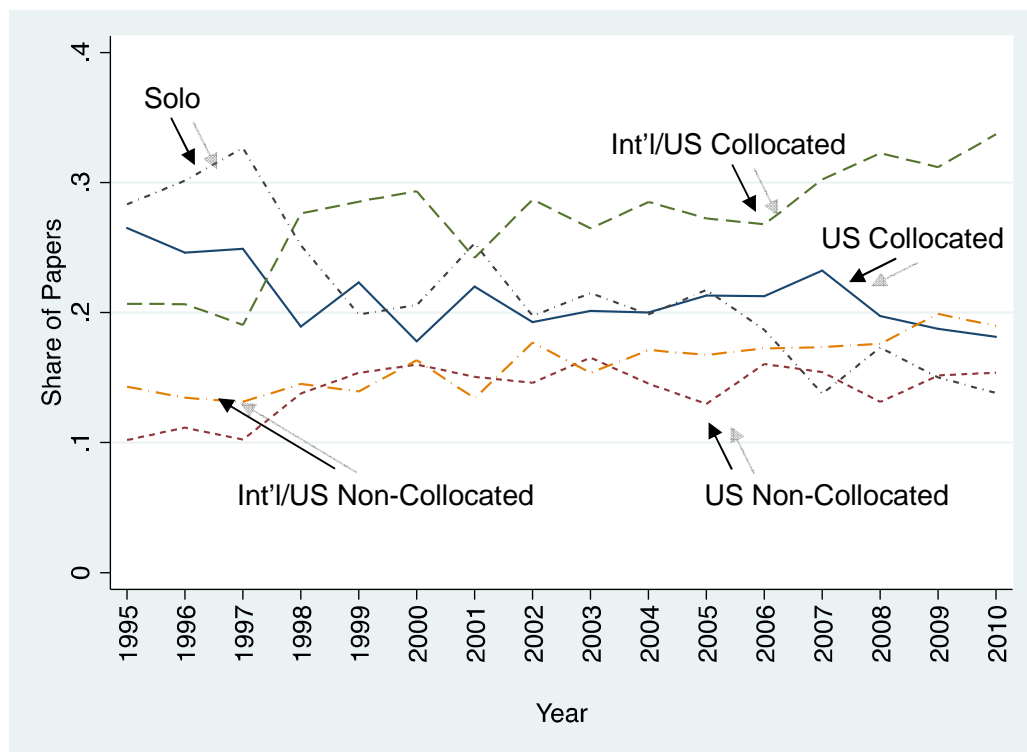


Figure 2b: Share of Papers by Collaboration Type, Nanotechnology

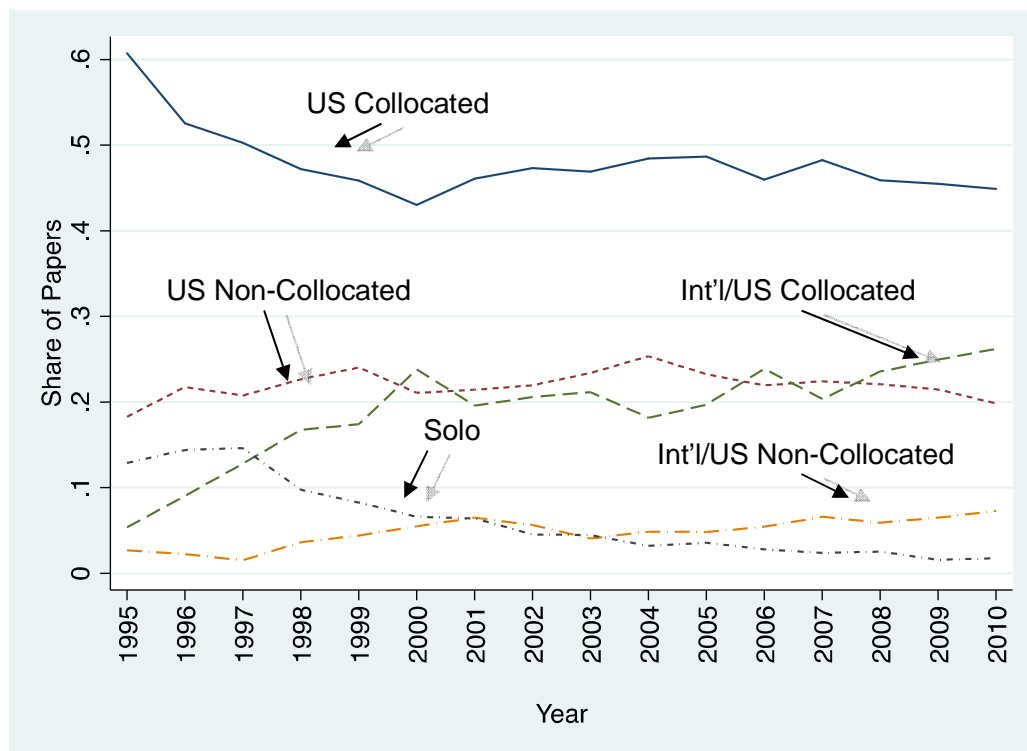


Figure 2c: Share of Papers by Collaboration Type, Biotechnology

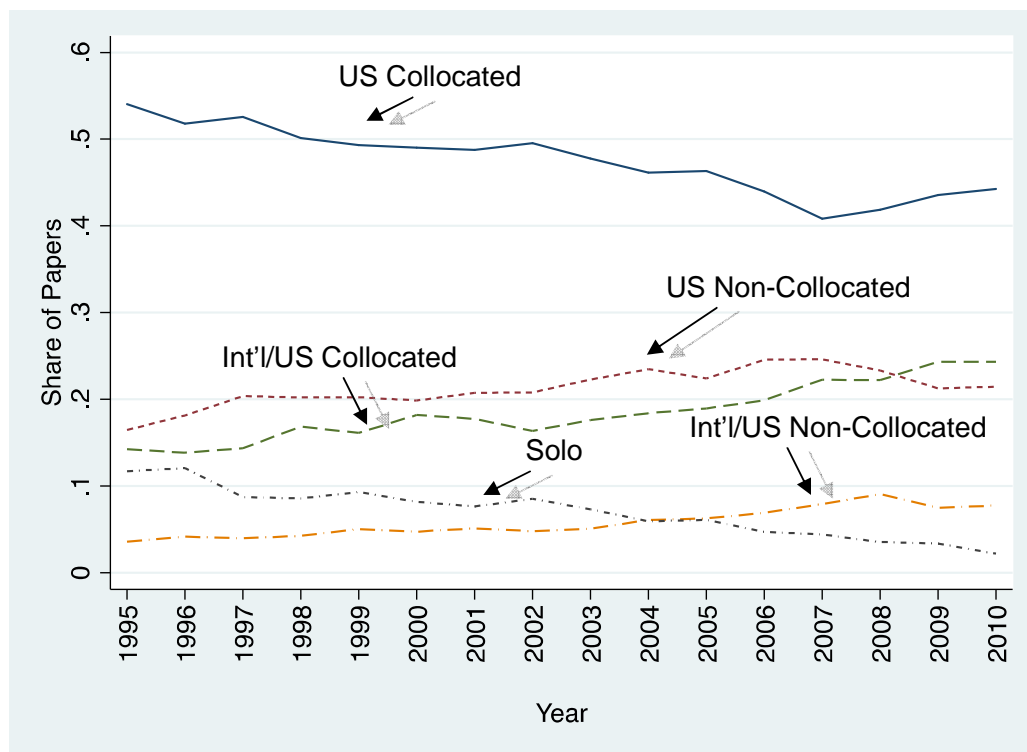
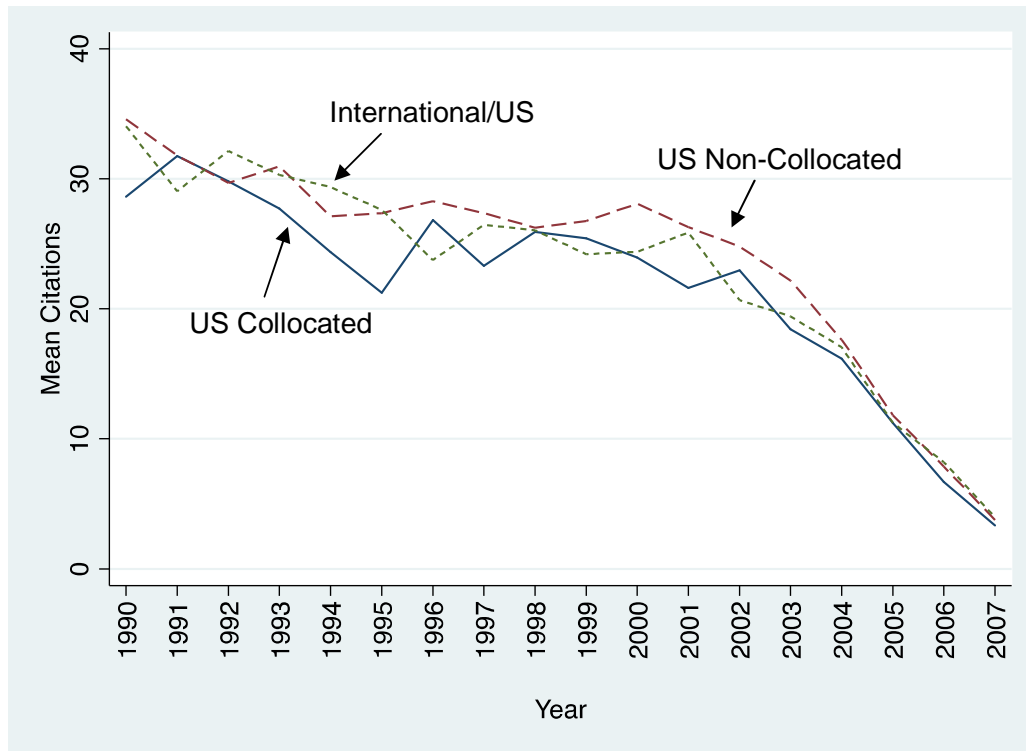
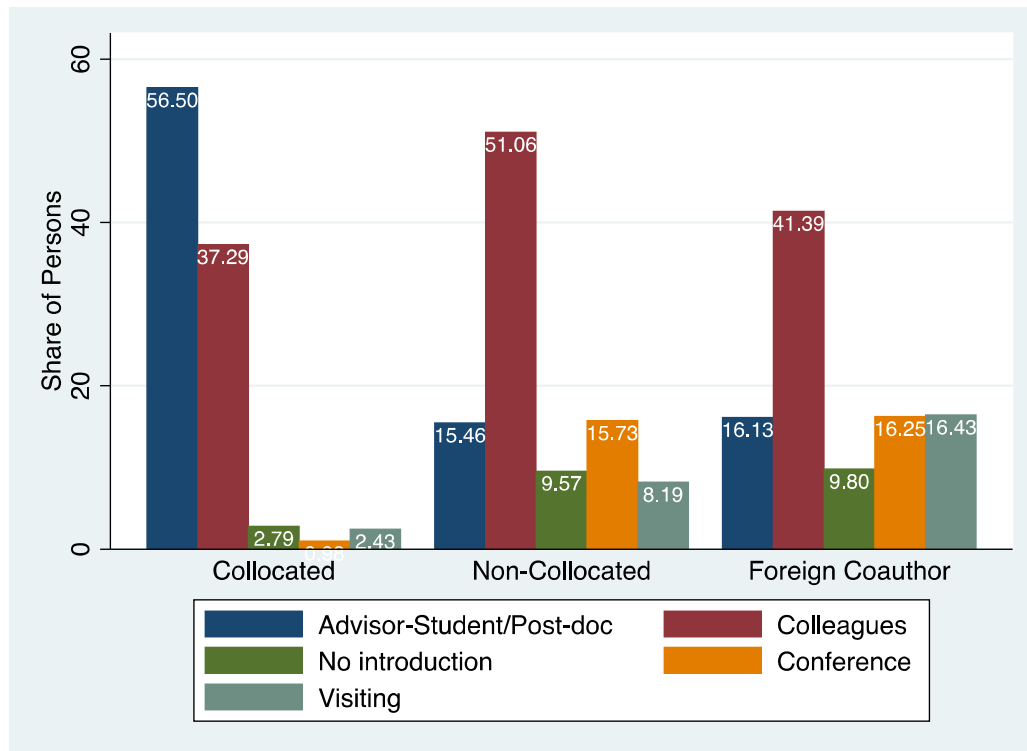


Figure 3. Citations By the Nature of Collaboration, All Fields by Year of Publication



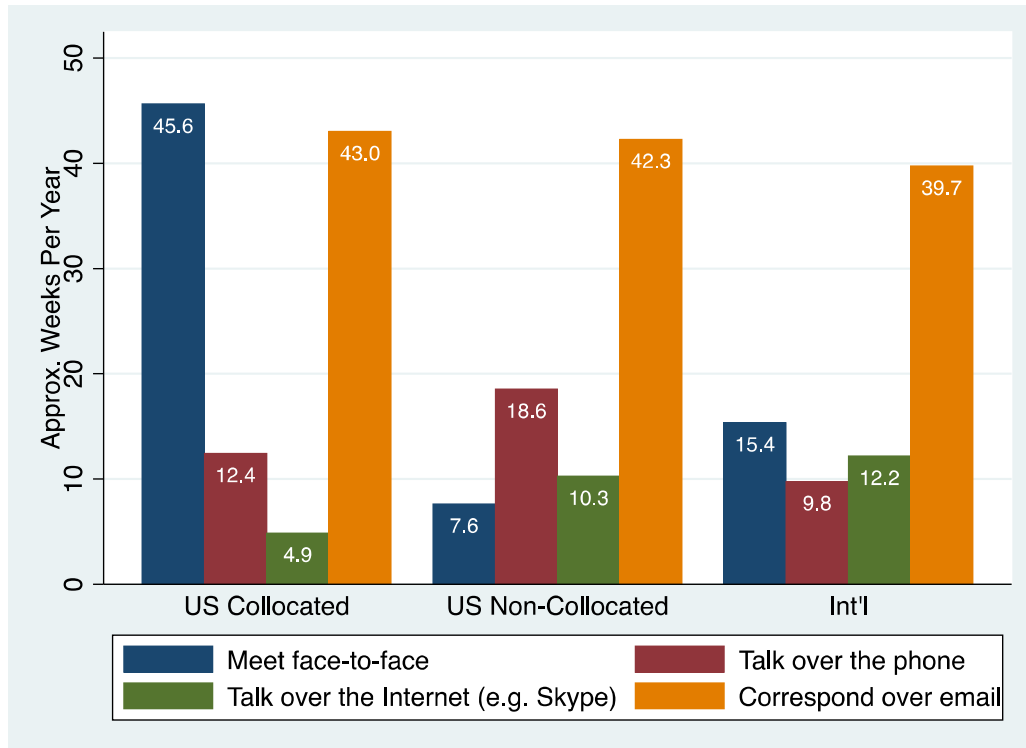
Notes: Figure shows forward citations of all papers in the Web of Science with at least one US author, and with journal subject categories of Physics, Particles & Fields; Nanoscience & Nanotechnology; and Biotechnology & Applied Microbiology, published from 1990-2007. Year indicates the year of publication of the cited paper.

Figure 4: Share of Persons Who Were First Met in a Given Way by the Nature of Collaboration



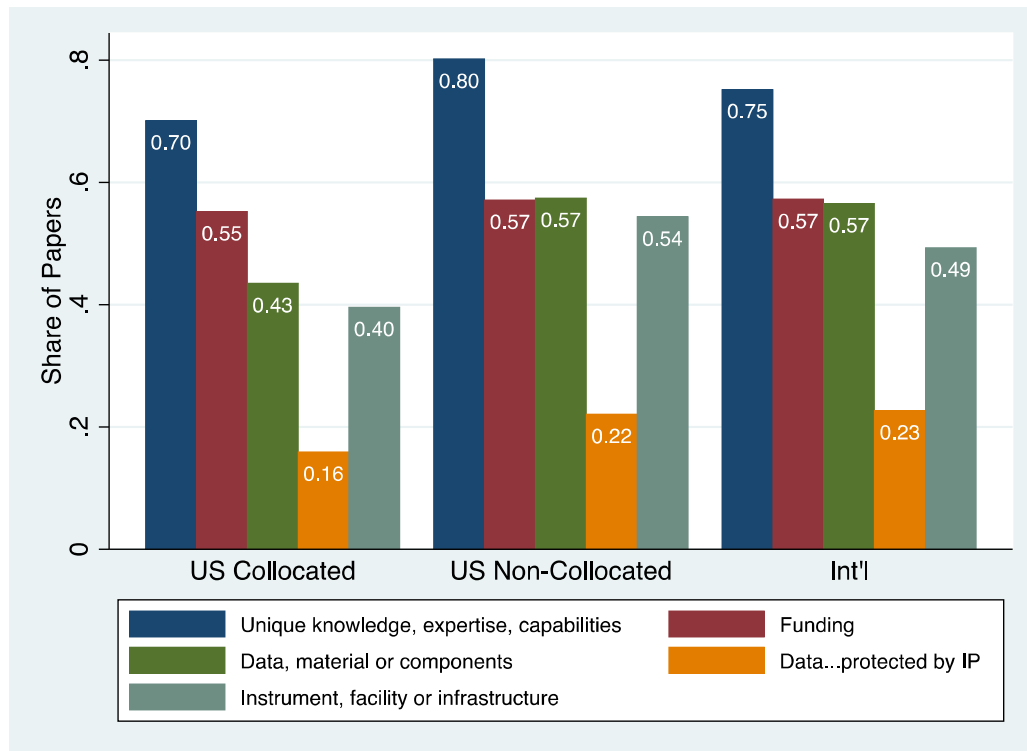
Notes: Share of all coauthors on papers for a given collaboration type. Question was phrased as “*How did you FIRST come in contact with each of these coauthors?*”

Figure 5. Overcoming Distance: Frequency of Communication Modes for 2-Author papers by the Nature of Collaboration (Approx. Weeks per Year)



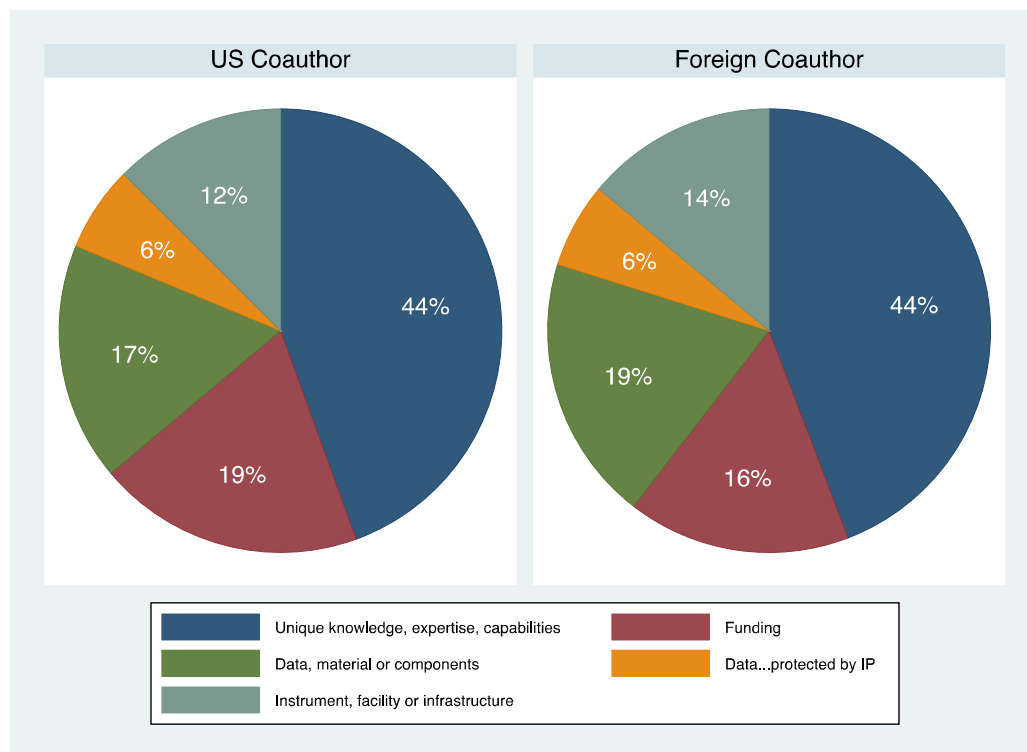
Notes: Question was phrased as “*When carrying out the research and writing for this article, how frequently did you use the following forms of communication with one or more of your coauthors?*” The possible choices were transformed into approximate number of weeks per year that each communication type was used: 6 = Every week (52), 5 = Almost every week (45), 4 = Once or twice a month (15), 3 = A few times per year (5), 2 = Less often than that (2), 1 = Never (0).

Figure 6: Contribution of Coauthors by the Nature of Collaboration



Notes: Share of papers for which the corresponding author reported at least one coauthor contributing the given resource. Question was phrased as “*Did any of the team members working on this article (including yourself) have access to one of the following resources that the other team members did NOT have, which made it important for you to all work together on this topic?*”

Figure 7. Contribution of US and Foreign Coauthors for 2-Author Papers



Notes: Share of US and foreign coauthors on 2-author papers only, as reported by the corresponding author.



## APPENDIX

Table A1. Team Size Summary Statistics

	Particle Physics	Nano	Biotech
Mean	21.86	4.56	4.74
Standard Deviation	82.49	2.49	3.61
Maximum	1062	32	202
<i>Percentiles</i>			
10th	1	2	2
50th	3	4	4
75th	4	6	6
95th	100	9	11
99th	523	13	16
N	40,474	31,934	68,731

Notes: Measures of number of authors on papers in the Web of Science published 1990-2010, with a US author (including solo author papers), and with journal subject categories of Physics, Particles & Fields; Nanoscience & Nanotechnology; and Biotechnology & Applied Microbiology.

Table A2. Optimal Team Size by Nature of Collaboration

	US Collocated	US Non-Collocated	All Int'l	Int'l with Cross-Country Funding
Yes	89.06	91.11	89.95	92.50
No, Additional	7.58	3.48	3.38	2.86
No, Fewer	3.37	5.40	6.67	4.64
N	1,663	574	1,154	280

Notes: Question phrased as “*Do you think that the size of your team was optimal?*” The cross-country funding question was phrased as “*In carrying out the research for this article, did any of the coauthors receive funding that was specifically aimed at supporting cross-country scientific collaboration?*”

Figure A1. Share of Coauthors Who Were First Met at a Conference

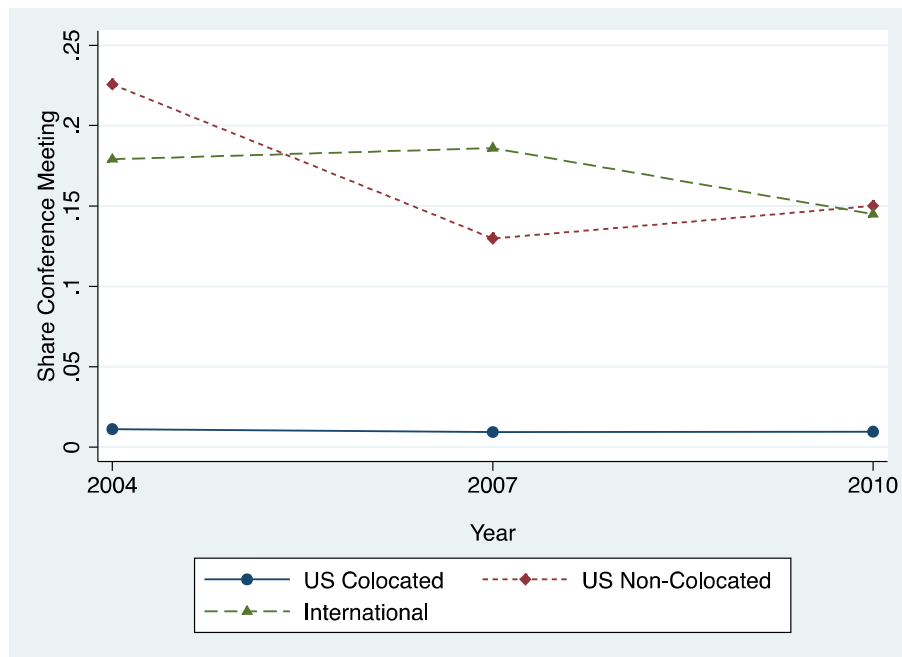


Figure A2. Share of Coauthors Who Were First Met as Advisor-Student/Postdoc

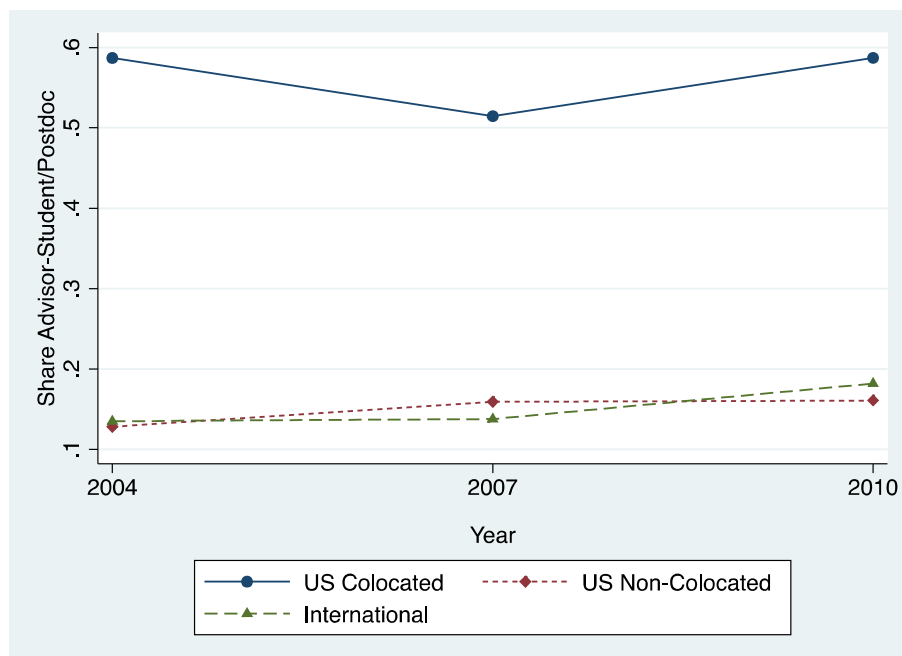


Figure A3. Share of Papers With a Coauthor Contributing Data, Material or Components

